

Angular cross section for $t\bar{t}$ production

The CDF Collaboration

Abstract

We measure the differential cross section for $t\bar{t}$ production as a function of the top production angle $\cos\theta_t$, employing a projection onto the Legendre polynomials to characterize the shape of the cross section. We observe agreement with the standard model prediction for all except the 1st Legendre moment. The top forward-backward asymmetry is dominated by the anomalously large 1st Legendre moment.

Contents

1	Introduction	2
1.1	Theoretical predictions	2
2	Event selection and modeling	3
3	Legendre moments	6
3.1	Theoretical predictions of moments	7
3.1.1	NLO (QCD+EWK) calculation	10
3.2	Estimating the moments	10
3.3	Correcting to parton level	13
4	Systematic uncertainties	14
5	Results	15
6	Conclusion	21
7	Acknowledgements	21

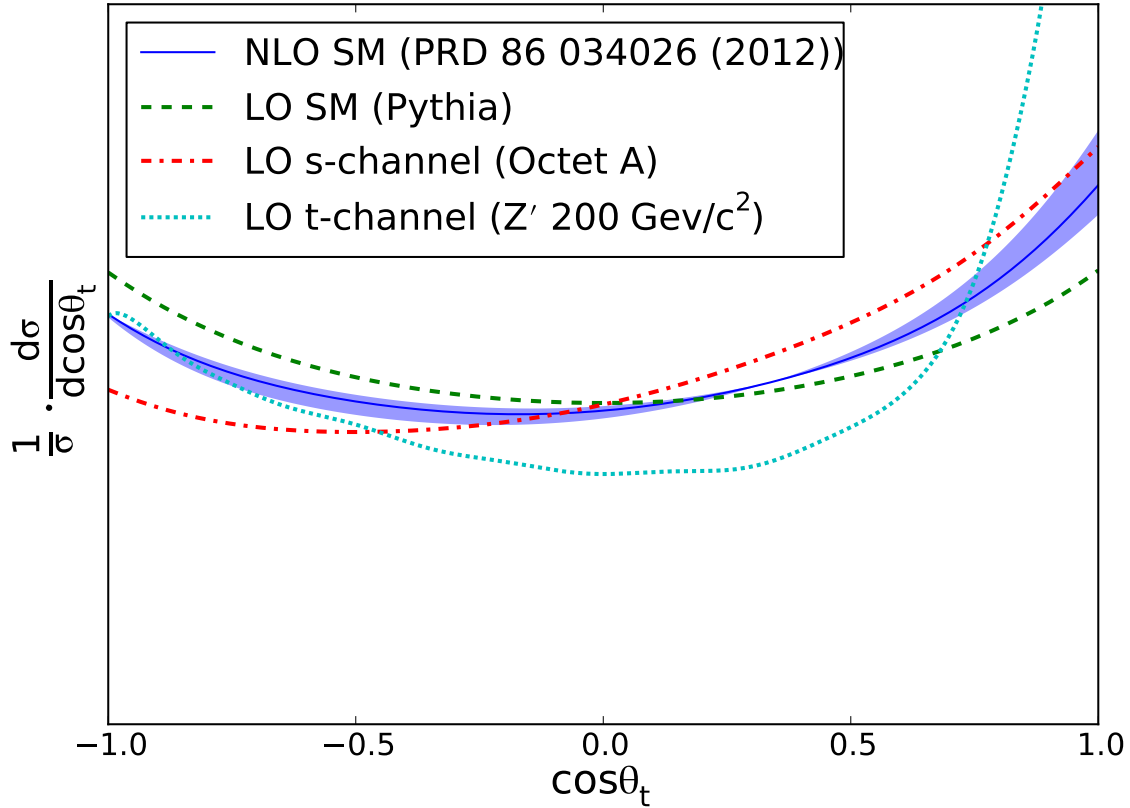


Figure 1: Distribution of $\cos \theta_t$ for various physics models

1 Introduction

Top pair production has been observed at CDF and D0 to exhibit a large forward-backward asymmetry (A_{FB}) that is not predicted by state-of-the-art Standard Model calculations (*Phys. Rev. D* **83** (112003) 2011 and *Phys. Rev. Lett.* **100** (142002) 2008). The asymmetry is a very simplistic statement about the distribution of the top production angle θ_t (the angle between the incoming proton momentum and the outgoing top momentum, as measured in the rest frame of the $t\bar{t}$ system) – equivalent to a 2-bin histogram. A more sophisticated approach is to investigate the full differential cross section as a function of $\cos \theta_t$, as this carries more information than A_{FB} alone. Additionally, the $\cos \theta_t$ differential cross section has the potential to discriminate among different physics scenarios.

1.1 Theoretical predictions

At leading order in the Standard Model, the process $q\bar{q} \rightarrow t\bar{t}$ proceeds via only a single s -channel diagram. The differential cross section has the general form $d\sigma/d\cos \theta_t = 1 + \cos^2 \theta_t$. When the process $gg \rightarrow t\bar{t}$ is included, the differential cross section becomes

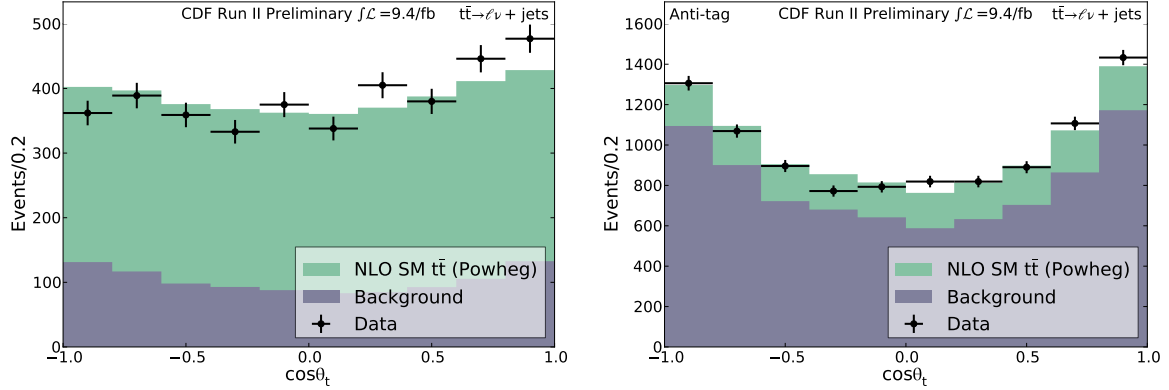


Figure 2: The distribution of $\cos \theta_t$ at the detector level in our data sample. The signal+background model generally performs well, except for the presence of a forward-backward asymmetry in the data that is not modeled by our signal MC. Also shown is the same distribution in the side-band with exactly zero b -tagged jets.

slightly more complicated. NLO contributions also add some perturbations to the tree-level prediction (see Figure 1).

There are also two broad classes of new physics which have been proposed to account for the observed A_{FB} . s -channel models contain a new axial, colored vector boson (axigluon) and produce the A_{FB} via interference effects between the parity-even (scattering via a SM gluon) and parity-odd (scattering via an axigluon) final states. t -channel models contain a new neutral flavor-changing boson which produces the A_{FB} by flavor-changing incoming up (or down) quarks into top quarks, which then scatter into the forward Rutherford peak. The s -channel models add a term to the SM that is linear in $\cos \theta_t$, whereas the t -channel models add a term that goes like $1/\hat{t} \sim 1/(1 - \cos \theta_t)$ (see *Phys. Rev. D* **83** (114027) 2011). Thus the s -channel models exhibit a deviation from the SM that is a deficit in the backward region that smoothly becomes an excess in the forward region. t -channel models produce a large excess in the very forward region, but relatively little deviation from the SM throughout the rest of the range of $\cos \theta_t$. See Figure 1 for LO and NLO SM calculations as well as benchmark models for both new physics scenarios.

2 Event selection and modeling

We utilize almost the same event selection, signal model, and background model as the lepton+jets A_{FB} measurement, documented in [arXiv:1211.1003](https://arxiv.org/abs/1211.1003). The only change in the event selection is to add a sample of events with three tight and at least one loose jet. The only change in the signal model is to use a new Powheg Monte Carlo sample that was generated using the luminosity profile of the entirety of Run II. From the full CDF Run II data set after data-quality requirements, totalling 9.4/fb, we require

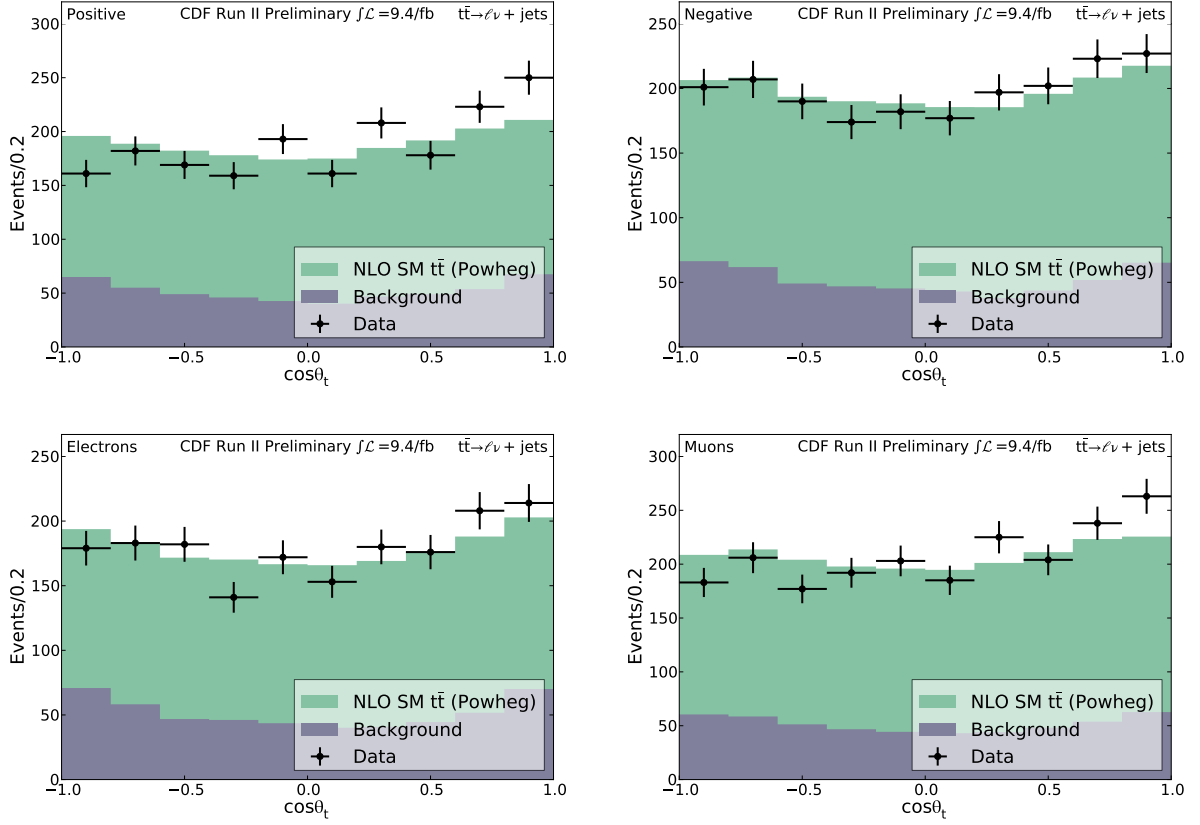


Figure 3: The distribution of $\cos \theta_t$ at the detector level separately for positive and negative leptons, and for electrons and muons. The data is well-modeled, and each subset of the data still displays a forward-backward asymmetry that is not present in the MC.

- exactly one well-reconstructed lepton with $p_T > 20 \text{ GeV}/c$,
- $\cancel{E}_T > 20 \text{ GeV}$,
- at least three tight jets with $E_T > 20 \text{ GeV}$ and $|\eta| < 2.0$,
- at least four total jets, where loose jets have a relaxed E_T requirement of only 12 GeV ,
- $H_T > 220 \text{ GeV}$,
- and at least one jet must be b -tagged.

Extensive validation of the data sample and signal and background model has been performed in the preceding analysis ([arXiv:1211.1003](https://arxiv.org/abs/1211.1003)), and we do not repeat it here. Because we are also studying a new variable ($\cos \theta_t$) that was not previously addressed, Figures 2 and 3 check the modeling of $\cos \theta_t$. We also check the modeling of $m_{t\bar{t}}$, the

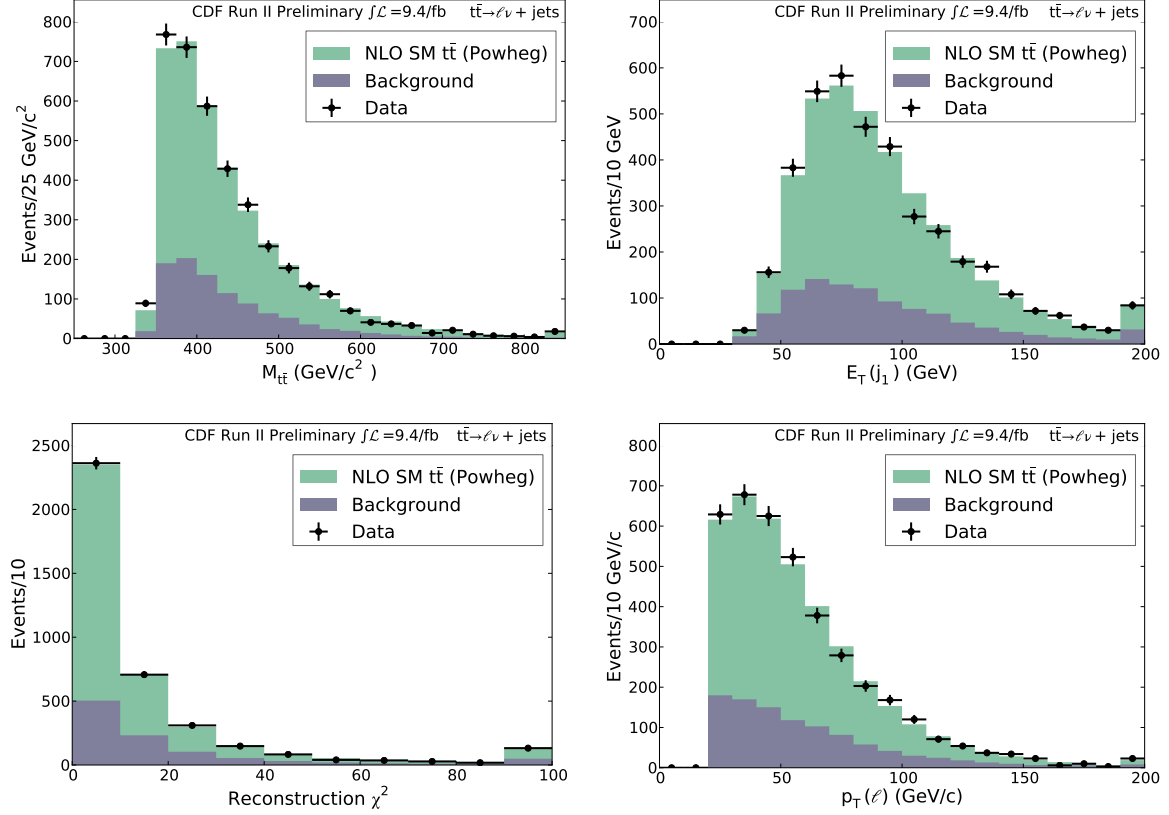
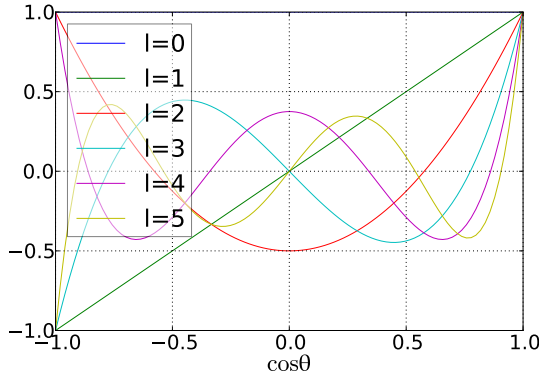


Figure 4: The distribution of various kinematic quantities. The modeling of our dataset is generally quite good.

CDF Run II preliminary 9.4/fb	≥ 4 jets
W +HF	481 ± 178
W +LF	201 ± 72
Non- W	207 ± 86
Single top	67 ± 6
Diboson	36 ± 4
Z +jets	34 ± 5
Total background	1026 ± 210
$t\bar{t}$ (7.4 pb)	2750 ± 427
Total prediction	3776 ± 476
Data	3864

Table 1: Predicted rates for the signal and the backgrounds, observed event count



ℓ	$P_\ell(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2 - 1)$
3	$\frac{1}{2}(5x^3 - 3x)$
4	$\frac{1}{8}(35x^4 - 30x^2 + 3)$
5	$\frac{1}{8}(63x^5 - 70x^3 + 15x)$

Figure 5: The shapes and formulas of the first six Legendre polynomials. The integral of every polynomial except for $\ell = 0$ is zero, and the even polynomials are symmetric while the odd polynomials are anti-symmetric.

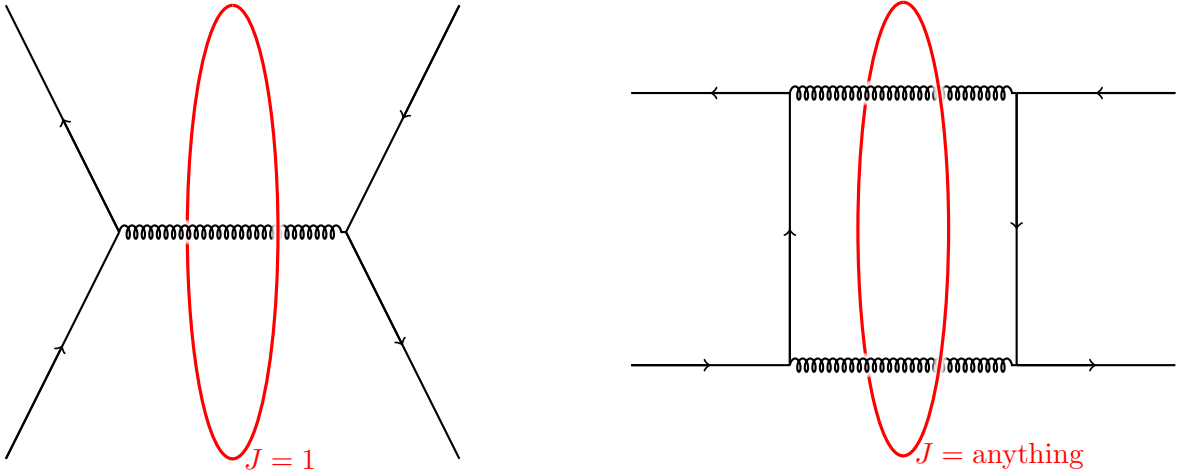


Figure 6: The Born diagram in the standard model proceeds only with a total angular momentum $J = 1$, while the box diagram, with a multi-particle intermediate state, may proceed with any angular momentum.

leading jet E_T , the transverse momentum of the charged lepton, and the kinematic reconstruction χ^2 in Figure 4. The number of events that we expect from signal and background and the number of events that we observe are given in table 1.

3 Legendre moments

The Legendre polynomials (and the spherical harmonics) appear frequently in discussion of angular distributions, and we will employ them in this analysis to describe the shape of the $\cos \theta_t$ distribution. This technique has also been employed in nuclear

and diffractive measurements in the past. Our measurement is of the first 9 Legendre moments (the coefficients of the first 9 Legendre polynomials), corrected to the parton level.

The first six Legendre polynomials are shown in Figure 5. The integral of all of the polynomials except for P_0 is zero. The even-degree polynomials are symmetric about $\cos\theta_t = 0$, while the odd-degree polynomials are anti-symmetric. Thus, non-zero Legendre moments of odd degree must be responsible for the A_{FB} . One of the main thrusts of this analysis is to determine what combination of odd-degree moments are responsible for the A_{FB} .

The 1959 paper by Jacob and Wick (*Ann. Phys. (NY)* **7** (4) 404-428) explores the general theory of $2 \rightarrow 2$ scattering of particles with mass and spin. For a general, unpolarized differential cross section, all of the angular dependence is entirely and very naturally described by the Legendre polynomials. The Legendre moments are determined by a sum over the helicities of the incoming and outgoing particles, two sums over total angular momentum states, the elements of the S matrix as a function of total angular momentum, helicity, and center of mass energy, and the Clebsch-Gordan coefficients.

3.1 Theoretical predictions of moments

This formalism allows us to connect the Legendre moments back to intuitively comprehensible physics. The only SM tree-level diagram for $q\bar{q} \rightarrow t\bar{t}$ is an s -channel diagram, and so only proceeds with a total angular momentum equal to the spin of the intermediate gluon, $J = s = 1$. Therefore, at tree-level in the SM, we should expect (via the Clebsch-Gordan coefficients) a contribution to the 0th, 1st, and 2nd Legendre moments ($|J_1 - J_2| \leq \ell \leq J_1 + J_2$). Since the gluon has even parity, and parity is conserved in SM QCD, we further expect that the 1st Legendre moment is zero, leaving us with non-zero moments at $\ell = 0$ and 2. In other, non- s -channel diagrams, any angular momentum is allowed, and so these diagrams, such as t -channel exchange or the QCD box diagram, produce contributions to all the Legendre moments from $\ell = 0$ to ∞ .

We explore four different models in this formalism. The predicted parton-level Legendre moments are shown in Figure 7. Because we are only interested in the shape of the differential cross section and not in the total cross section, we have normalized all of the predictions so that the zeroth moment, $a_0 = 1$.

The first model is the LO SM QCD calculation of Pythia. In this model, as discussed, we observe a large a_0 and a_2 , but a_1 is zero. We also see some small contribution to a_4 , a_6 , and a_8 . This is due to the inclusion of the $gg \rightarrow t\bar{t}$ process, which includes a t -channel diagram.

Second is a NLO SM calculation. This is a dedicated computation performed by Bernreuther and Si (*Phys. Rev. D* **86** (034026) 2012 and private communication), which includes the full Standard Model (QCD, QED, weak) at next-to-leading order. As the box diagrams are now included, we pick up a correction to all the moments,

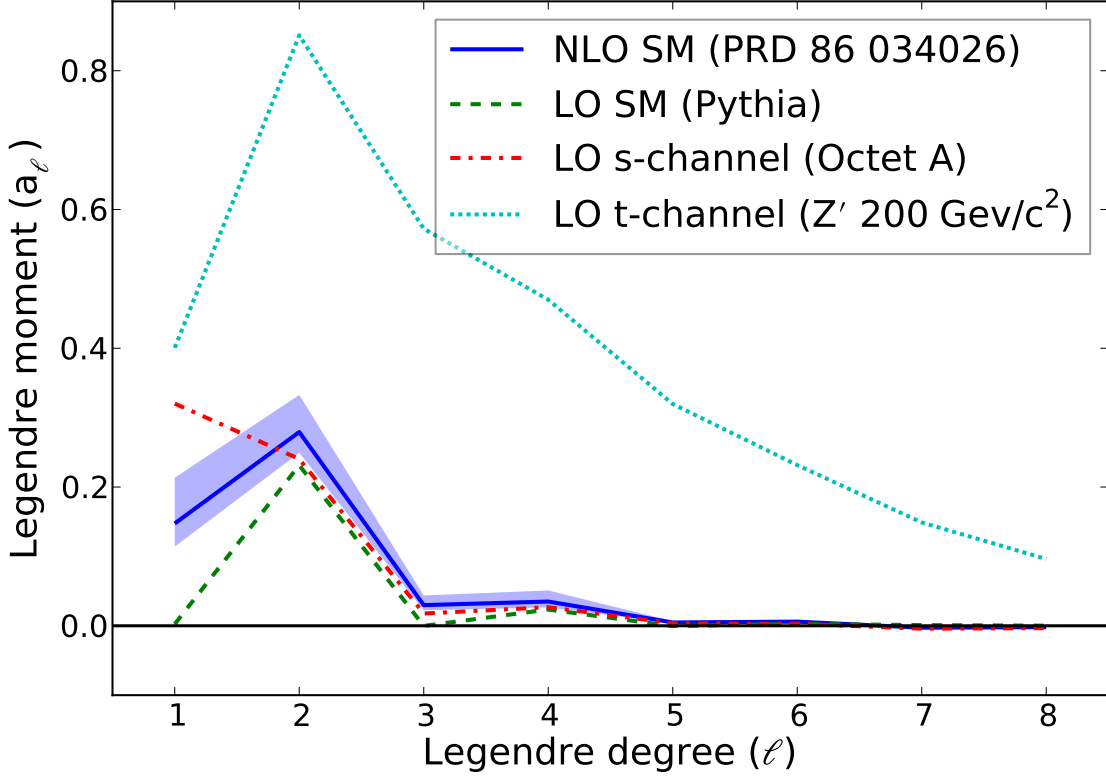


Figure 7: The Legendre moments for several physics scenarios.

including non-zero contributions to the odd moments. Details on this model are ahead in section 3.1.1.

The s -channel model we use is generated by MadGraph at tree-level, with an axigluon added ($m_{G'} = 2 \text{ TeV}/c^2$ and coupling strength to quarks $g = 3/2$, where the coupling to top quarks is opposite the coupling to all other quarks). As expected, this adds a linear term to the Pythia calculation, which becomes a deviation in the 1st moment alone. This model has been explored thoroughly in previous analyses (called “Octet A”, see [arXiv:1211.1003](https://arxiv.org/abs/1211.1003)), and does a very good job of reproducing all kinds of kinematic distributions, as well as reproducing the total inclusive A_{FB} . Octet A was in fact specifically tuned to reproduce the inclusive A_{FB} in the 5/fb analysis without adversely affecting the distribution of $m_{t\bar{t}}$ or the total cross section. It is very representative of s -channel models with a high-mass axigluon in the general case.

The t -channel model is also generated by MadGraph. It includes a flavor-changing Z' ($m_{Z'} = 200 \text{ GeV}/c^2$) with a $u-t-Z'$ coupling ($g = 0.65$). As the added diagram is not an s -channel diagram, the angular momentum arguments cannot simply be applied, and we expect a very large contribution to all the Legendre moments under study. While it is not completely clear whether this model is representative of t -channel models in

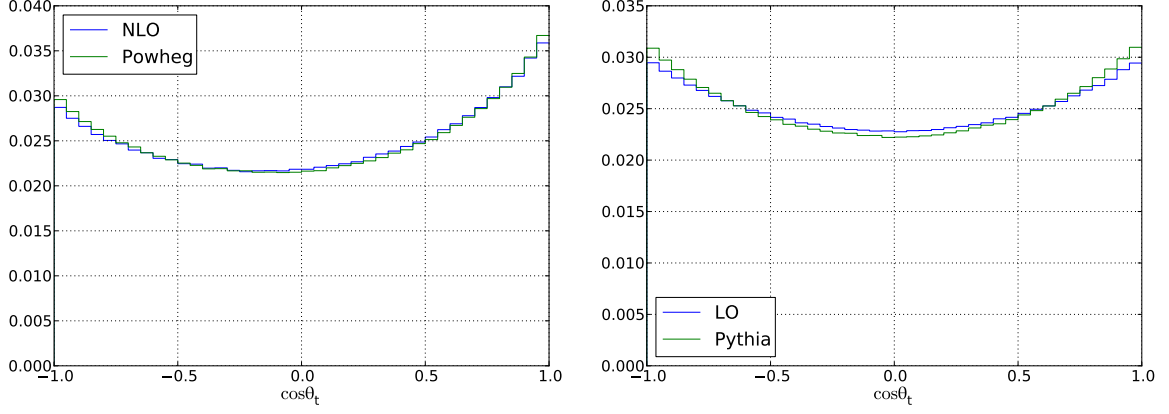


Figure 8: Results of dedicated $\cos\theta_t$ calculation compared to Pythia and to Powheg.

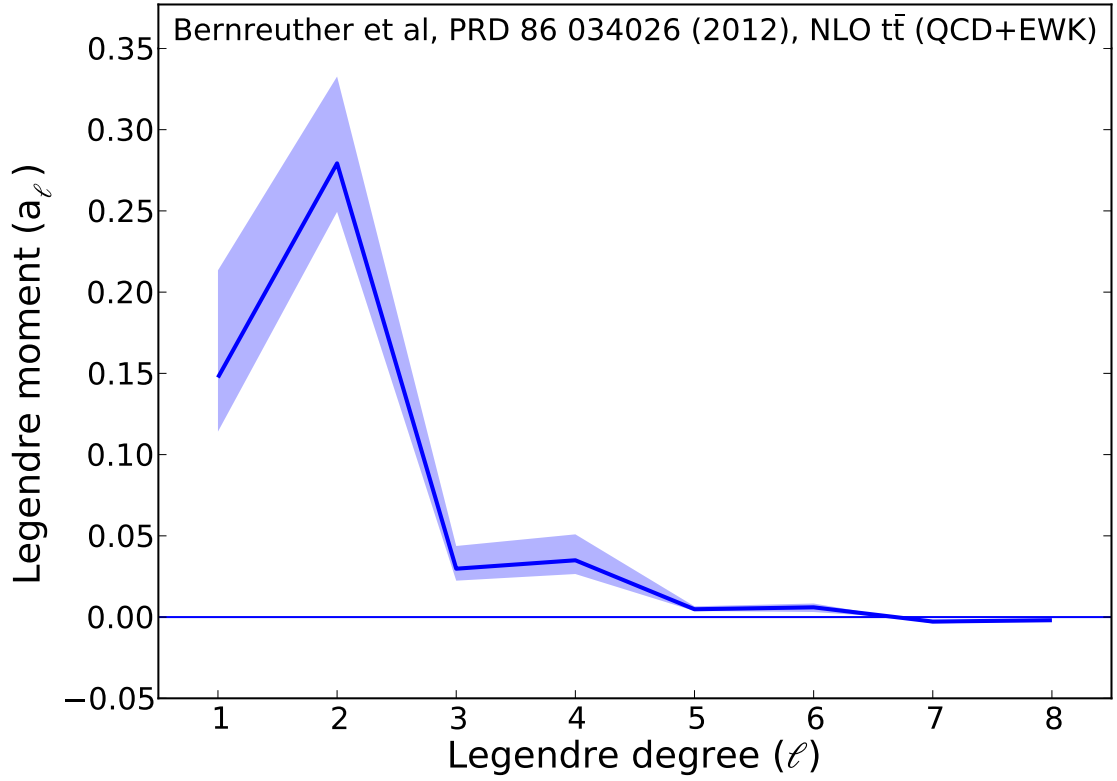


Figure 9: Legendre moments of the dedicated $\cos\theta_t$ calculation.

the general case, it is worth noting that t -channel models are already disfavored by the LHC and Tevatron data.

3.1.1 NLO (QCD+EWK) calculation

Various papers by Hollik and Pagani ([arXiv:1107.2606](#)), Kuhn and Rodrigo ([arXiv:1109.6830](#)), and Manohar and Trott ([arXiv:1201.3926](#)) have pointed out that electroweak effects have a substantial impact on predictions of A_{FB} . By taking NLO QCD diagrams and replacing various combinations of gluon lines with photon, Z , and W lines, we find additional contributions to the total cross section, as well as alterations to various differential cross sections. These effects have been calculated to increase A_{FB} by approximately 26% by altering the Δy differential cross section used in previous A_{FB} measurements. In order to perform a differential cross section measurement, it is important to know how the corrections affect the predicted shape of the differential cross section.

Werner Bernreuther and Zong-Guo Si (*Phys. Rev. D* **86** (034026) 2012 and private communication) have performed a dedicated calculation of the $\cos \theta_t$ differential cross section when considering only leading order in QCD, next-to-leading order in QCD, NLO in QCD plus corrections from weak bosons, and NLO in QCD+EWK (the full Standard Model). The LO part of this calculation compares well to Pythia, and the NLO (QCD only) part compares well to Powheg (Figure 8). There is some minor disagreement in that our Monte Carlos sit in excess of Bernreuther’s calculations at large $|\cos \theta_t|$. We attribute this to the presence of Pythia parton showering (which rewrites the momenta of the showered partons to conserve 4-momentum), which is not present in Bernreuther’s calculations. The Legendre moments of Bernreuther’s calculations are shown in Figure 9. The principal change between NLO QCD and NLO QCD+EWK is to the first moment. This calculation, and these moments, are the benchmark we use throughout the analysis as the Standard Model.

3.2 Estimating the moments

In order to avoid discarding information by coarsely binning our data in a histogram, we use our data as if it were a histogram with an infinite number of bins. This is the “empirical measure” of mathematical statistics. In a histogram with an infinite number of bins, each bin is filled at most once, and almost all bins are not filled at all. For those bins that are filled, the width of the bin is zero, but the integral over the bin is one, so the height of the bin is infinite. This is therefore equivalent to viewing the data as a sum of Dirac delta functions, where there is one delta function at each data point. This approach also greatly simplifies the problem of estimating the Legendre moments that correctly describe the distribution of our data.

In order to compute the Legendre moments, a_ℓ of a function $f(\cos \theta)$, one exploits the orthogonality of the Legendre polynomials,

$$a_\ell = \frac{2\ell + 1}{2} \int_{-1}^1 d\cos \theta f(\cos \theta) P_\ell(\cos \theta). \quad (1)$$

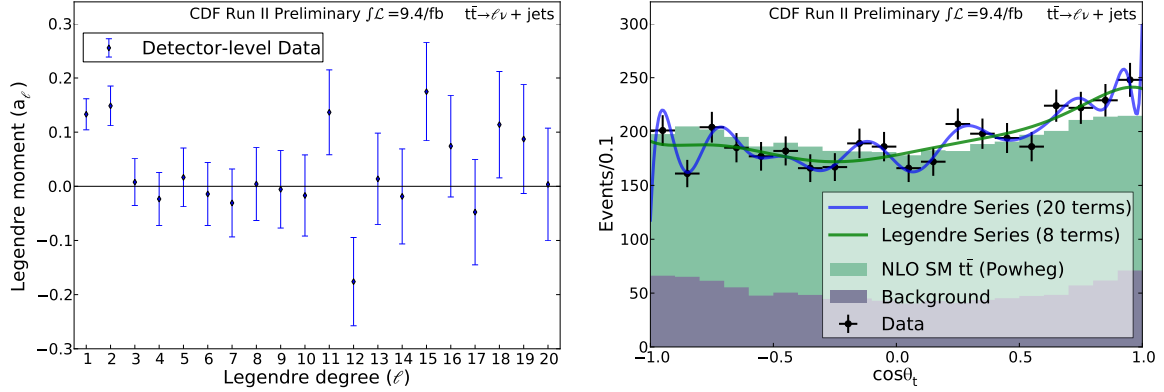


Figure 10: The estimated Legendre moments at the detector-level in the CDF data before background subtraction, and the resulting Legendre series superimposed on a histogram of the data. The uncertainties on the moments derive from the diagonal of the covariance matrix (Equation (6)).

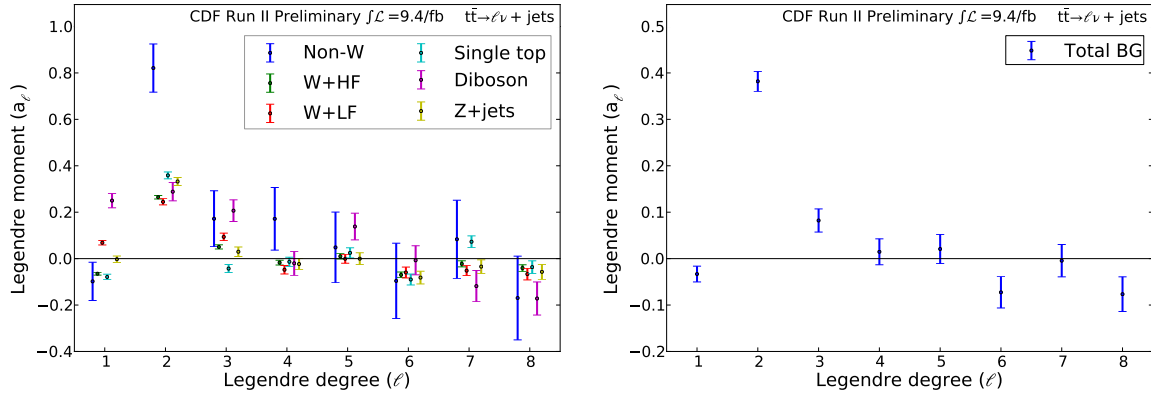


Figure 11: The estimated Legendre moments at the detector-level arising from various backgrounds mis-reconstructed as $t\bar{t}$ events.

When the function $f(\cos \theta)$ is a sum of delta functions,

$$f(\cos \theta) = \sum_i \delta(\cos \theta - \cos \theta_i), \quad (2)$$

this integration becomes trivial:

$$a_\ell = \frac{2\ell + 1}{2} \int_{-1}^1 d\cos \theta \sum_i \delta(\cos \theta - \cos \theta_i) P_\ell(\cos \theta) \quad (3)$$

$$= \frac{2\ell + 1}{2} \sum_i \int_{-1}^1 d\cos \theta \delta(\cos \theta - \cos \theta_i) P_\ell(\cos \theta) \quad (4)$$

$$= \frac{2\ell + 1}{2} \sum_i P_\ell(\cos \theta_i), \quad (5)$$

where the sum is over observed values of $\cos \theta$. In words, we look at each event in our data (indexed by i), and we calculate the value of $\cos \theta_i$ for that event. We then evaluate each Legendre polynomial at that value of $\cos \theta_i$. Then we sum over all the events in our data to obtain an estimate for the coefficient (Legendre moment) of each Legendre polynomial.

This estimate technically follows a “compound Poisson” distribution, which is the distribution that results from summing a Poisson-distributed number of weights, where the weights follow their own (unspecified) distribution. In our case, a single weight corresponds to a single event, and so a weight is a vector of Legendre polynomial output values, $P_\ell(\cos \theta_i)$. We sum these vector-valued weights over the set of events, and the number of events is, of course, Poisson distributed, so we do indeed have a compound Poisson distribution.

We forgo a rigorous exposition of the parameters of this distribution in favor of a simpler heuristic argument. When we count events, we obtain an estimate of the total event rate that follows a Poisson distribution. The best estimate of the rate (the mean of the underlying Poisson distribution) is just the number of events, and the best estimate of the uncertainty on that rate (the standard deviation of the underlying Poisson distribution) is the square root of the number of events. As long as the rate is large enough, the mean and uncertainty are sufficient to describe the distribution because we are in the regime where a Gaussian approximation is valid.

We can recast this counting experiment in the language of a compound Poisson distribution. In that case, the weights are scalars and are always one. The best estimate of the mean of the underlying compound Poisson distribution is the sum of the weights, and the best estimate of the standard deviation is the square root of the sum of the squares of the weights, a familiar formula. This pattern of the mean being estimated by the sum of weights and the variance being estimated by the sum of squares of weights can be extended to the case where the weights are not always one, and also to the case where the weights are vectors rather than scalars. The only change needed to use vector-valued weights is that the square should be replaced by the tensor product, which yields a covariance matrix rather than a scalar variance. The Gaussian approximation regime is also then replaced by the multivariate-Gaussian approximation regime.

In this analysis, we do have sufficient statistics that we are well into the Gaussian approximation regime, and so the distribution of our estimate of the Legendre moments is approximately multivariate Gaussian described by a mean as given above from the sum of weights and a covariance matrix derived from the sum of (tensor product) squares of weights:

$$\Sigma_{\ell\ell'} = \sum_i P_\ell(\cos \theta_i) \cdot P_{\ell'}(\cos \theta_i). \quad (6)$$

This method is used to estimate the Legendre moments for the detector-level measurement of $\cos \theta_t$ for the distribution of the events that pass our event selection criteria. We obtain the distribution of the moments shown in Figure 10. As shown in Figure 10, the statistical noise in the data disproportionately affects the estimation of the

higher-order moments. Conversely, the higher-order moments express the statistical fluctuations (local over- and under-densities).

We follow the same method to estimate the Legendre moments of the background distributions, and then subtract those from the Legendre moments in the data. The data and background covariances are added together to obtain the background-subtracted estimate and covariance. The moments describing the backgrounds, both individual backgrounds and the total background, are shown in Figure 11.

3.3 Correcting to parton level

In order to have a useful measurement that can be directly compared to theory and to other experiments, we must correct our results for the effects of mismeasurement, misreconstruction, and inefficiencies. In general, the detector-level differential cross section can be related to the parton-level differential cross section by a convolution,

$$f_{\text{det}}(\cos \theta_{\text{det}}) = \int_{-1}^1 d \cos \theta_{\text{part}} K(\cos \theta_{\text{det}}, \cos \theta_{\text{part}}) f_{\text{part}}(\cos \theta_{\text{part}}), \quad (7)$$

where K is a transfer function that expresses the probability that an event with a given θ_{part} will be observed with a given θ_{det} . Naturally if the integral of K over detector-level values,

$$\int_{-1}^1 d \cos \theta_{\text{det}} K(\cos \theta_{\text{det}}, \cos \theta_{\text{part}}) = Pr(\cos \theta_{\text{part}}), \quad (8)$$

is less than one, that expresses the probability that such a parton level event will be observed at all.

In order to estimate the parton-level result, we must estimate the transfer function K and then deconvolve the detector-level result. We estimate the transfer function using the fully-simulated MC, in which we know both θ_{det} and θ_{part} for every selected event, and we know whether each generated event was accepted, allowing us to estimate both detector smearing and acceptance effects. So, using the same “empirical measure” trick as before,

$$K(\cos \theta_{\text{det}}, \cos \theta_{\text{part}}) \equiv \frac{1}{f_{\text{gen}}(\cos \theta_{\text{part}})} \sum_i \delta(\cos \theta_{\text{det}} - \cos \theta_{\text{det},i}) \cdot \delta(\cos \theta_{\text{part}} - \cos \theta_{\text{part},i}), \quad (9)$$

where $f_{\text{gen}}(\cos \theta_{\text{part}})$ is the generator-level distribution before any acceptance effects.

In general, deconvolution of continuous functions is intractable. However, by exploiting the Legendre polynomials, we are able to discretize this particular deconvolution. After converting the convolution into the Legendre moment basis, we have

$$a_{\ell}^{\text{det}} = \frac{2\ell+1}{2} \sum_m K_{\ell m} a_m^{\text{part}}, \quad (10)$$

which we can easily solve with simple matrix inversion. The detector-level covariance, Equation (6) is propagated through the matrix multiplication in the usual fashion, by

multiplying the unfold matrix to both left and right of the covariance matrix. The resulting statistical error estimation is tested using pseudoexperiment pulls, which follow a χ^2 distribution with the expected number of degrees of freedom. The χ^2 is actually a “Mahalanobis distance”, which is a χ^2 in the presence of non-zero correlations,

$$\chi^2 = (a_\ell - \hat{a}_\ell)C^{-1}(a_\ell - \hat{a}_\ell), \quad (11)$$

where a_ℓ is the expected distribution, \hat{a}_ℓ is the distribution resulting from the pseudoexperiment, and C^{-1} is the inverse of the covariance matrix estimated from the pseudoexperiment.

In most analyses involving an unfold to the parton level, this discretization is accomplished by binning the data and the transfer function into histograms. The resulting transfer matrix is typically numerically close to a singular matrix and therefore inversion leads to grossly magnified error bars. In order to control this problem, regularization is introduced via the assumption that the correct parton-level differential cross section is smooth (as in *Nucl. Instrum. Meth.* **A372** (469-481) 1996). However, in the case of the Legendre moment unfold, the transfer matrix is well-conditioned, and simple inversion is sufficient. We do not apply any regularization in our Legendre moment unfold.

We also note that these techniques can be applied to a similar measurement at the LHC, substituting the direction of the longitudinal boost of the $t\bar{t}$ system for the proton direction.

4 Systematic uncertainties

We consider two categories of systematic uncertainty. Background systematics affect the rate or the shape of the background predictions. Signal systematics affect the unfolding matrix. In order to evaluate the effect of each source of systematic uncertainty, we first alter either the background prediction or the unfolding matrix, and then we re-do the background subtraction and unfold. This provides us with a new set of measured, parton-level Legendre moments, a_ℓ^{varied} . We compare these varied moments to an appropriate nominal set of moments a_ℓ^{nominal} , and obtain a vector of shifts in the moments due to that source of systematic uncertainty,

$$\delta_\ell = a_\ell^{\text{varied}} - a_\ell^{\text{nominal}}. \quad (12)$$

These shifts represent an uncertainty on the unfolded moments which has 100 % correlation across all the moments. As a result, we produce a covariance matrix with 100 % correlation by taking the tensor product of the vector of shifts with itself,

$$\Sigma_{\ell m} = \delta_\ell \cdot \delta_m. \quad (13)$$

In this fashion, we make covariance matrices representing each of the sources of systematic uncertainty, and sum them all up with the statistical-only covariance matrix

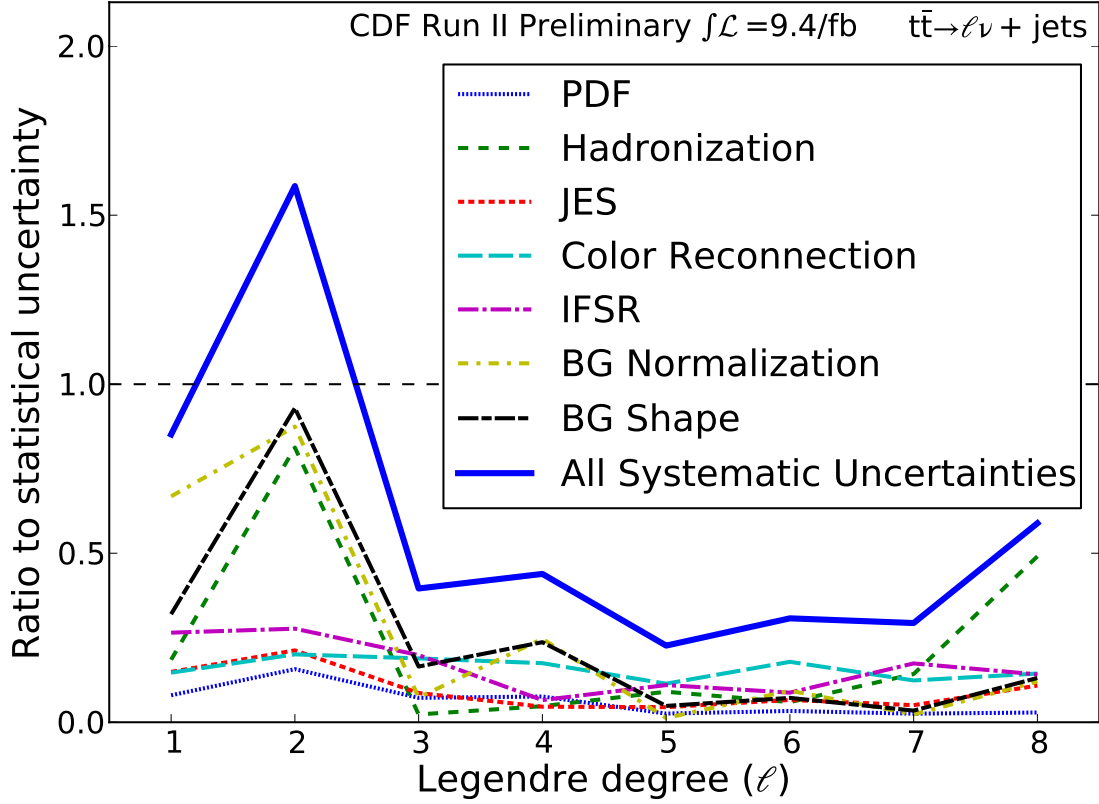


Figure 12: The size of each systematic, and of the total systematic uncertainty, relative to the statistical uncertainty.

to produce a full covariance matrix completely describing the uncertainty on our measurement.

We consider systematic uncertainties due to the uncertainty on the predicted background rates and shapes, on jet energy measurements, in the modeling of hadronization, color reconnection, and initial- and final-state radiation, and on the parton distribution functions of the proton. The sum of all the systematics is shown in Figure 12. We are statistics-limited in our measurement of all of the Legendre moments except for a_2 .

5 Results

We observe (Figure 13 and Table 2) that the Legendre moments agree well with the NLO SM prediction except for the 1st moment. The first moment is anomalously large. Because no other odd moments are anomalously large, the first moment is the primary contributor to the A_{FB} . While the SM calculation is still in flux as various theorists are working on new calculations, the 1st moment currently sits 2.1σ from the central NLO SM prediction (Figure 15).

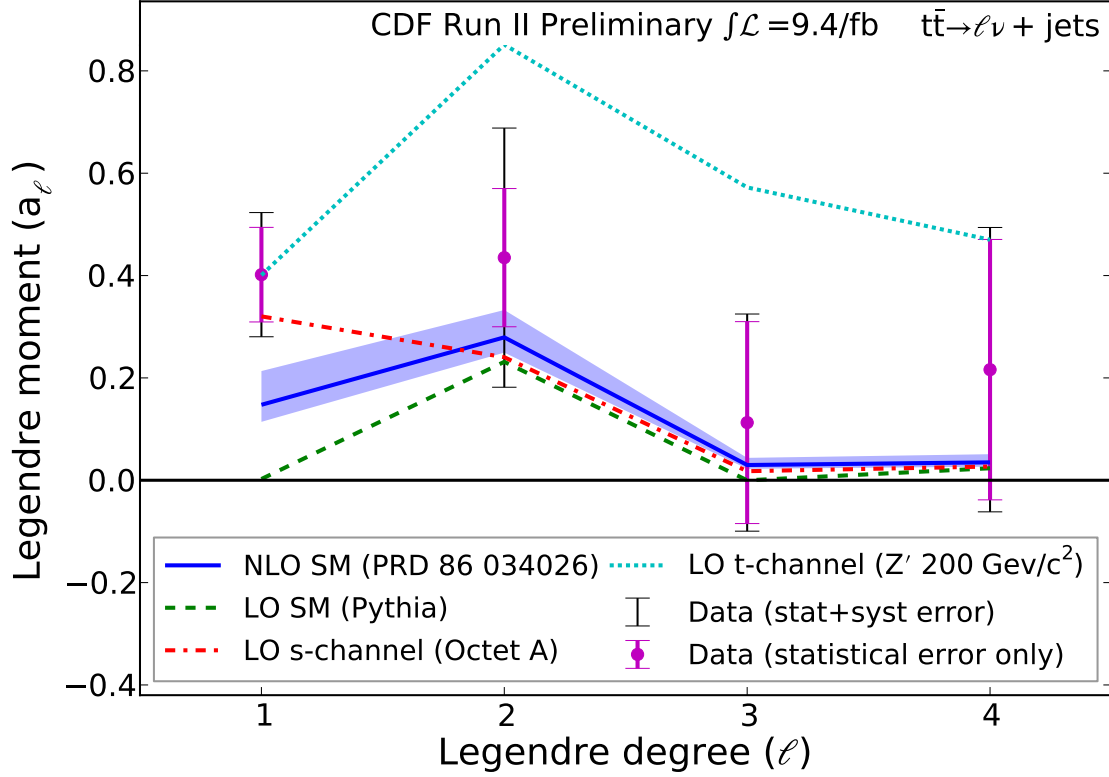


Figure 13: Measured Legendre moments in CDF Run II data, with statistical and systematic uncertainties, compared to the predictions of four different models. Uncertainties are highly correlated. The band around the NLO SM prediction is the result of varying the scale from $m_t/2$ to $2m_t$, with m_t as the central prediction.

While previous analyses measured A_{FB} , which is a blanket statement about the totality of the anti-symmetric portions of the differential cross section, this is the first analysis to break down those anti-symmetric terms into some sort of component parts, which allows us to make much stronger statements about the underlying physics. Additionally, this is the first analysis to measure the *symmetric* portions of the differential cross section at all. This measurement will help to constrain both models of new physics and new calculations of the Standard Model, and will hopefully help to push the state of the art in high-scale QCD calculations forward.

Figure 16 shows that the A_{FB} is dominated by the first moment. The other moments contribute only negligibly to the A_{FB} . The A_{FB} produced by the non-linear terms is totally consistent with the SM prediction, but of course, the A_{FB} produced by the linear term is well in excess of the SM.

Since these observations are fully corrected to the parton level, they are directly comparable to theory and to other experiments. The proper way to compare them is via the Mahalanobis distance discussed above. This simple test statistic can be computed

CDF Run II Preliminary $\int \mathcal{L} = 9.4/\text{fb}$		
	Data (stat+syst)	NLO (QCD+EWK)
Legendre degree (l)	Legendre moment (a_ℓ)	Legendre moment (a_ℓ)
1	$0.40 \pm 0.09 \pm 0.08$	$0.15^{+0.066}_{-0.033}$
2	$0.44 \pm 0.14 \pm 0.21$	$0.28^{+0.053}_{-0.030}$
3	$0.11 \pm 0.20 \pm 0.08$	$0.030^{+0.014}_{-0.007}$
4	$0.22 \pm 0.25 \pm 0.11$	$0.035^{+0.016}_{-0.008}$
5	$0.11 \pm 0.32 \pm 0.07$	$0.0048^{+0.002}_{-0.001}$
6	$0.24 \pm 0.39 \pm 0.12$	$0.0060^{+0.002}_{-0.003}$
7	$-0.15 \pm 0.46 \pm 0.14$	$-0.0028^{+0.001}_{-0.001}$
8	$0.16 \pm 0.56 \pm 0.33$	$-0.0019^{+0.0003}_{-0.0003}$

Table 2: Measured Legendre moments in CDF Run II data, with statistical and systematic uncertainties (which are highly correlated).

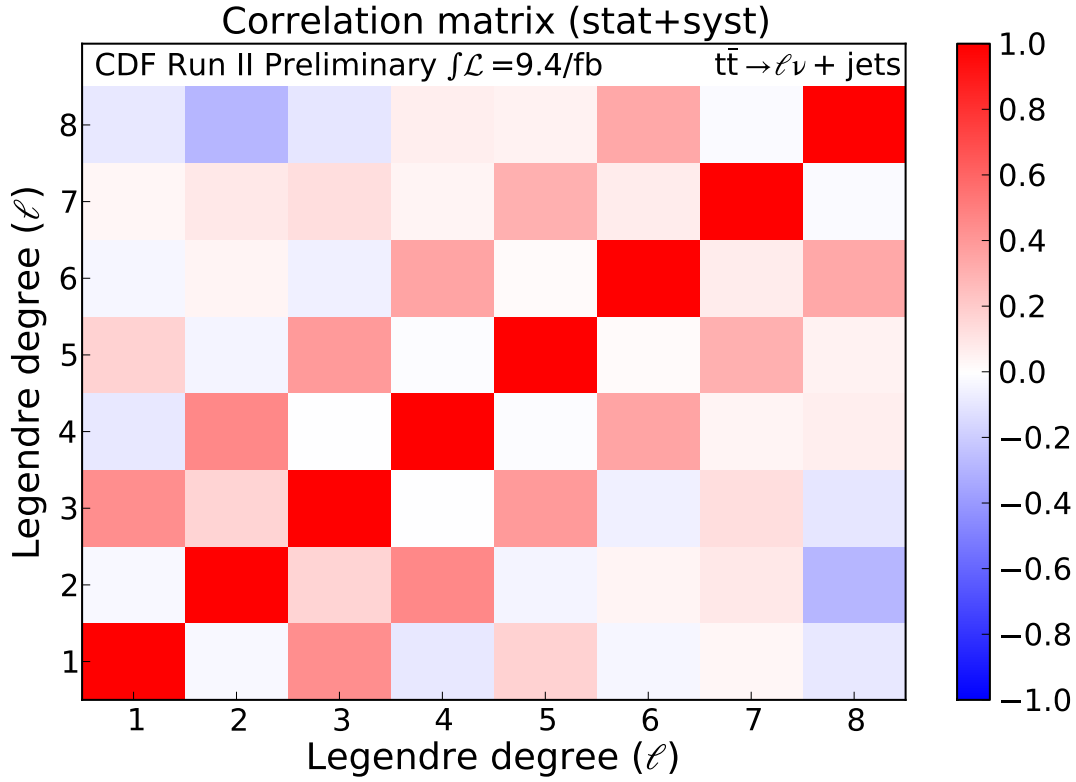


Figure 14: The correlation matrix among the measured moments. Note that each moment is principally correlated with its next-to-nearest neighbor.

CDF Run II Preliminary $\int \mathcal{L} = 9.4/\text{fb}$

	1	2	3	4
1	1.47×10^{-2}	-7.60×10^{-4}	1.14×10^{-2}	-2.95×10^{-3}
2	-7.60×10^{-4}	6.41×10^{-2}	8.96×10^{-3}	3.29×10^{-2}
3	1.14×10^{-2}	8.96×10^{-3}	4.50×10^{-2}	-8.18×10^{-5}
4	-2.95×10^{-3}	3.29×10^{-2}	-8.18×10^{-5}	7.72×10^{-2}
5	6.86×10^{-3}	-3.70×10^{-3}	2.72×10^{-2}	-7.32×10^{-4}
6	-1.73×10^{-3}	4.12×10^{-3}	-4.88×10^{-3}	4.00×10^{-2}
7	2.01×10^{-3}	1.05×10^{-2}	1.36×10^{-2}	5.49×10^{-3}
8	-7.34×10^{-3}	-4.69×10^{-2}	-1.33×10^{-2}	1.13×10^{-2}
	5	6	7	8
1	6.86×10^{-3}	-1.73×10^{-3}	2.01×10^{-3}	-7.34×10^{-3}
2	-3.70×10^{-3}	4.12×10^{-3}	1.05×10^{-2}	-4.69×10^{-2}
3	2.72×10^{-2}	-4.88×10^{-3}	1.36×10^{-2}	-1.33×10^{-2}
4	-7.32×10^{-4}	4.00×10^{-2}	5.49×10^{-3}	1.13×10^{-2}
5	1.06×10^{-1}	2.31×10^{-3}	4.85×10^{-2}	1.13×10^{-2}
6	2.31×10^{-3}	1.63×10^{-1}	1.42×10^{-2}	8.81×10^{-2}
7	4.85×10^{-2}	1.42×10^{-2}	2.32×10^{-1}	-5.45×10^{-3}
8	1.13×10^{-2}	8.81×10^{-2}	-5.45×10^{-3}	4.17×10^{-1}

Table 3: The covariance matrix describing the uncertainty on the measured parton-level moments. This matrix may be inverted to calculate a χ^2 statistic with 8 degrees of freedom that will allow future theoretical calculations to be compared directly to this measurement.

using only the predicted moments, the observed moments, and the estimated covariance matrix (Table 3). It follows a χ^2 distribution with 8 degrees of freedom.

One application of this Mahalanobis distance is to make assumptions about the values of some of the moments. After noting that a_1 is the only moment in disagreement with the SM prediction, we can obtain a more precise, but more model-dependent, measurement of a_1 by explicitly assuming that a_2 through a_8 are as predicted by the Standard Model within the scale variation uncertainties. We combine this assumption with our measurement by minimizing a joint χ^2 (the BLUE method). This gives a best-fit value

$$a_1 = 0.390 \pm 0.108, \quad (14)$$

which constitutes a reduction in uncertainty of about 10 %, and a shift in the central value of only 2.5 %.

The axigluon models, as noted, differ from the Standard Model principally by the addition of a term which is linear in $\cos \theta_t$. The coefficient of that term is determined by a combination of the axigluon mass, width, and couplings. As equation (14) is likely the

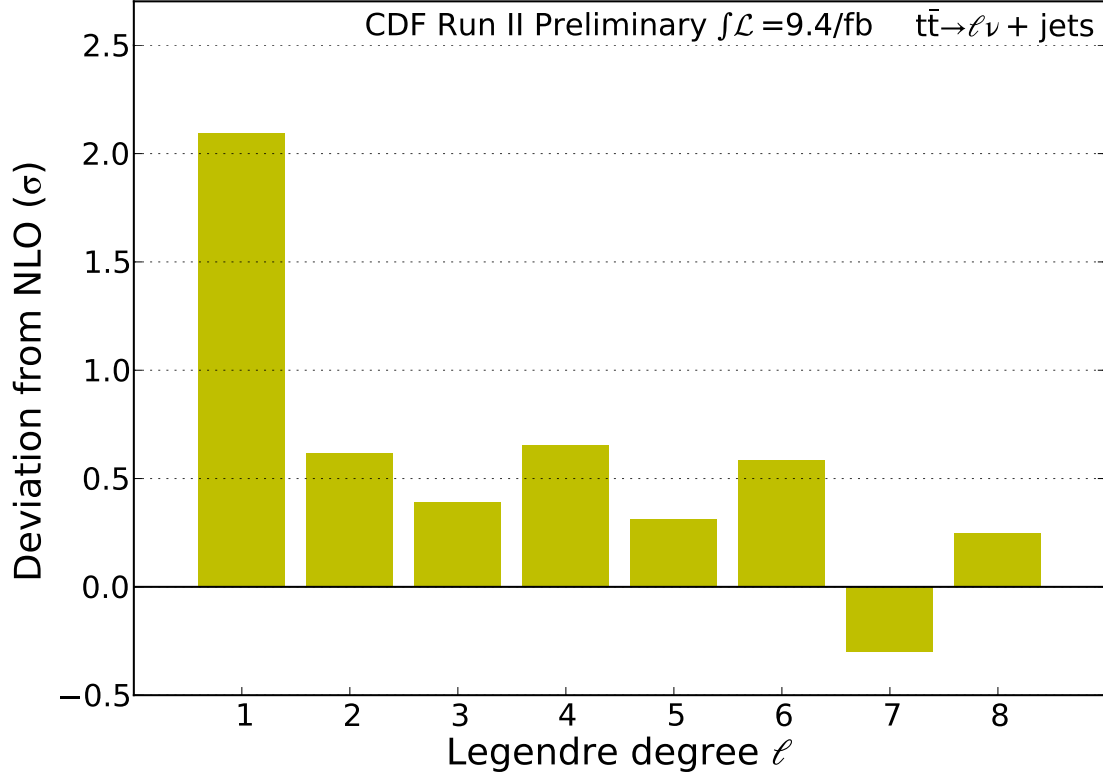


Figure 15: Deviation of the observed Legendre moments from the NLO SM prediction with scale m_t , divided by the measurement uncertainty on each moment.

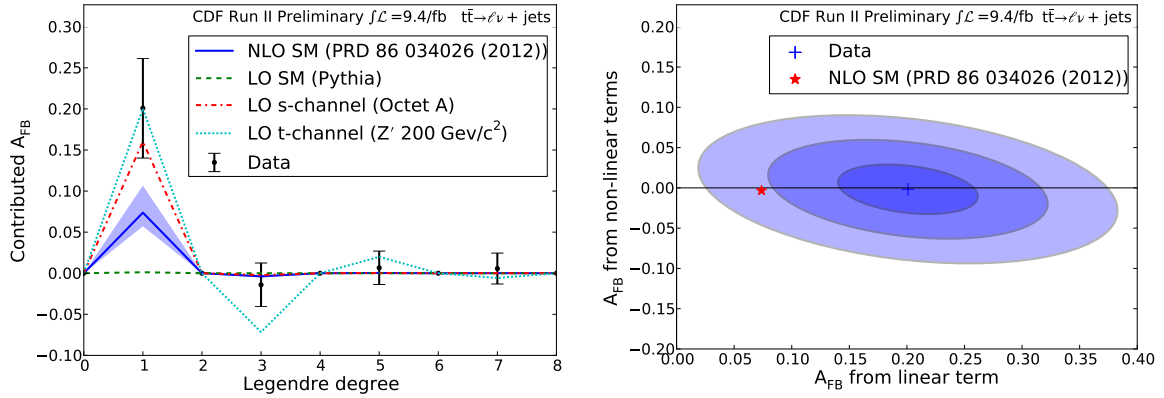


Figure 16: The contribution of each moment to the A_{FB} , and the A_{FB} resulting from the linear term (first moment) and all the other, non-linear, terms combined.

best measurement currently possible of the linear term in the differential cross section, it could be used to estimate the parameters of an axigluon model. One could take the mass-dependent measurement of A_{FB} from [arXiv:1211.1003](https://arxiv.org/abs/1211.1003) and assume that, as

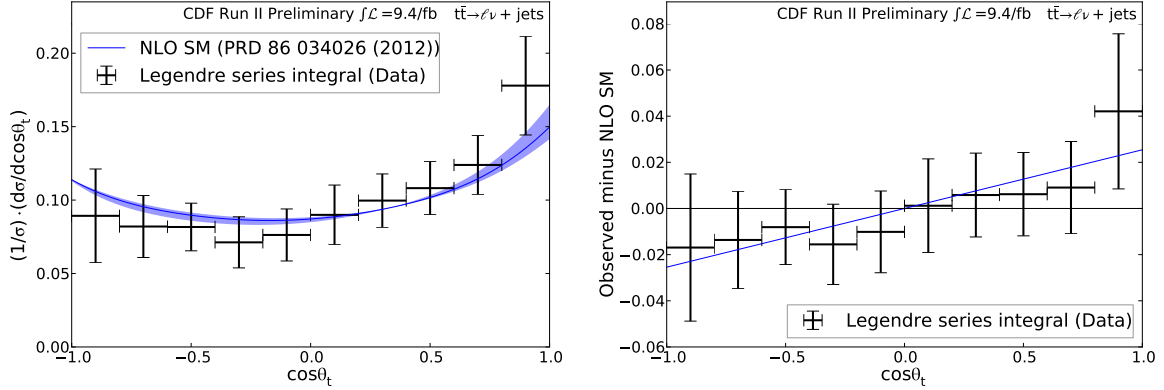


Figure 17: Integrating the Legendre series over finite-width bins gives the fraction of cross section in each bin of $\cos \theta_t$. Uncertainties are highly correlated and are dominated by the large uncertainties on the higher-order moments.

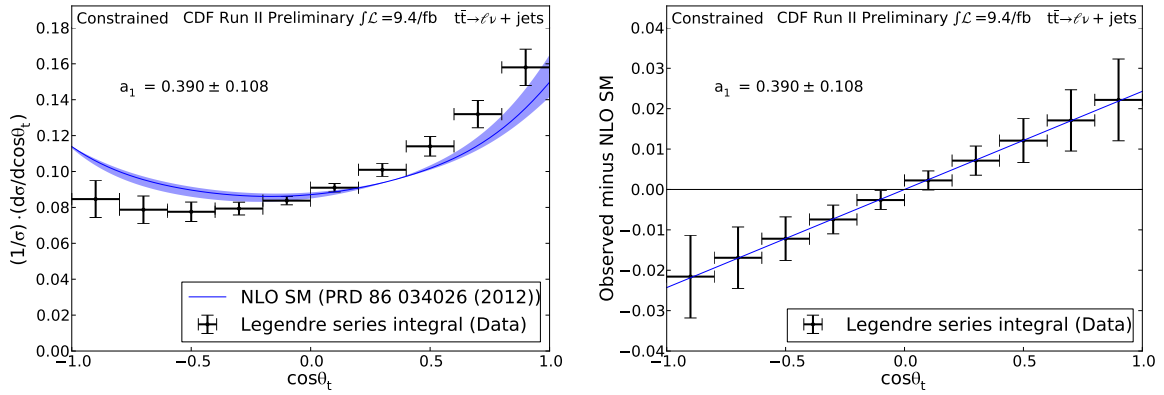


Figure 18: We introduce the assumption that the NLO SM correctly predicts moments $a_2 - a_8$, and integrate the resulting Legendre series over bins again.

in the inclusive case, the A_{FB} was entirely due to a linear term, and therefore have a mass-dependent linear coefficient, which should be sufficient to indirectly determine what axigluon masses, widths, and couplings are allowed and/or favored by the CDF data.

We also integrate the Legendre series resulting from the observed Legendre moments over the width of several finite-sized bins in $\cos \theta_t$ both before (Figure 17) and after (Figure 18) introducing the additional assumptions leading to equation (14). Since we are not attempting to measure the total cross section at all, we leave the integral scaled to unity, so this gives the fraction of measured cross section falling into each bin in $\cos \theta_t$. We chose 10 bins because we measure 8 moments, and 10 is a nice, round number. Of course, the uncertainties are still highly correlated, and the size of the uncertainties is largely due to the large uncertainty on the higher-order Legendre moments.

6 Conclusion

We have measured the differential cross section as a function of the production angle, $\cos\theta_t$, using the coefficients of the Legendre polynomials to describe its shape. We observe an anomalously large first moment, $a_1 = 0.40 \pm 0.09(\text{stat}) \pm 0.08(\text{syst})$. This corresponds to an anomalously large term in the differential cross section. All other measured moments are in good agreement with the prediction of the Standard Model at next-to-leading order. This excess first moment is sufficient to explain the excess forward-backward asymmetry; other measured moments contribute only negligibly to the A_{FB} .

7 Acknowledgements

We thank W. Bernreuther, Z.-G. Si, S. Jung, and T. Tait for their assistance in the preparation of our theoretical models.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; the Australian Research Council (ARC); and the EU community Marie Curie Fellowship contract 302103.